RBB Economics

Annexure B: Response to the PHMI's Provisional Findings

Supplier Induced Demand

RBB Economics, 15 October 2018

Executive Summary

The Private Healthcare Market Inquiry ("PHMI") released its Provisional Findings and Recommendations Report ("the PF") on 5 July 2018. The PF includes an analysis of supplier induced demand ("SID"). This report, prepared by RBB Economics LLP ("RBB"), provides an assessment of this analysis, focussing in particular on the extent to which it relates to private healthcare facilities.

The PF finds that the increases in healthcare expenditures that it has observed are, to a large extent, driven by increased utilisation of private healthcare services. It considers that a major driver of increased utilisation is SID on the part of both healthcare practitioners and facilities. However, the analysis contained within the PF does not provide a sound basis to conclude that, to the extent that SID exists, it is in any way a consequence of conduct by private hospitals.

First, the existing literature on SID of healthcare services recognises that healthcare facilities do not act as an agent in the provision of healthcare services, and hence cannot act directly to induce demand. As a consequence, it is practitioners, not facilities, that are typically placed at the centre of theories of SID.

Nonetheless, the PF argues that healthcare facilities stimulate SID indirectly via arrangements with practitioners. Specifically, the PF alleges that hospitals provide practitioners with incentives to increase overall utilisation rates (i.e. to induce demand) through contractual relationships and offering equity shareholdings in hospitals.

However, the PF's analysis of relationships between practitioners and hospitals does not support this conclusion. In particular, the contracting arrangements between facilities and practitioners discussed in the PF do not incentivise practitioners to increase the absolute level of admissions (which would amount to SID).

On the contrary, such arrangements only incentivise practitioners to channel a given portion of whatever the total number of admissions that they do make to the hospital in question. In addition, as we explain in more detail below, practitioners' shareholdings in healthcare facilities do not provide a material incentive to practitioners to induce demand. Accordingly, the PF does not present any basis on which to conclude that the existence and extent of SID in South Africa can be attributed to private healthcare facilities.

Second, the PF's comparison of hospital utilisation rates in South Africa with utilisation rates in other OECD countries, from which it concludes that "*South Africa tends to over-service"*, is fundamentally flawed, and therefore unreliable.[1] In this regard, the PF compares utilisation rates in South Africa with those of 17 OECD comparator countries for all hospital admissions and selected "discretionary" procedures. The PF finds that:

1 The PF, paragraph 429, page 246.

- overall hospital utilisation rates were higher in South Africa than in all but 2 OECD comparator countries, after adjusting for differences in the age distribution of country's populations;
- age-adjusted admissions rates were higher in South Africa than the average of all comparator countries for 6 out of 7 "discretionary" procedures; and
- ICU admissions rates were higher in South Africa than in 8 comparator countries.

The PF considers these findings to be evidence of the presence of SID in South Africa.

However, the PF's analysis does not, in fact, show that South African utilisation rates are systematically higher than those of the OECD comparator countries. In particular, overall age-adjusted admissions in South Africa are only marginally higher than the OECD comparator country average rate, and age-adjusted utilisation rates in South Africa are lower or not materially different from the average OECD comparator country rate for 4 out of the 7 "discretionary" procedures considered in the PF's analysis.

Moreover, there are significant issues with the comparability of OECD comparator countries with South Africa, such that any differences between countries cannot be reliably ascribed to the existence of SID. For example, a large proportion of the comparator counties with lower utilisation rates than South Africa have significant resource constraints, which may therefore be expected to naturally result in lower levels of utilisation than in South Africa. In contrast, South Africa's utilisation rates tend to be similar, or lower, than the rates in countries that, like South Africa's private healthcare system, are unlikely to be resource constrained.

By the same token, despite acknowledging the importance of accounting for a wide range of demand-side factors that are likely to drive cross-country differences in utilisation rates, the PF only accounts for a single factor, namely differences in the age distribution between South Africa and the OECD comparator countries. This provides a further reason why any differences between countries cannot be reliably ascribed to the existence of SID.

Third, the PF's finding that "*rates of hospital admissions are positively associated with levels of supply of hospital beds",* which it finds, in turn, to be supportive of SID acting through hospital capacity, is based on an analysis that contains a number of material flaws.[2] When these flaws are addressed, the analysis yields results that show no such relationship. Indeed, even if the results presented in the PF were taken as given they would only provide very weak support for the PF's finding.

Specifically, the PF presents a series of regression analyses that seek to determine, at the level of overall admissions, and also for particular specialties, whether there exists a positive relationship between the supply of practitioners and/or hospital beds and levels of utilisation (after controlling for relevant explanatory factors). These results indicate that positive relationships exist between both of these variables and admission levels at a local level.

---

[2] The PF, paragraph 427, page 245.

However, is should be noted from the outset that the results from the PF's model of overall hospital admissions shows that there is, at best, a very weak relationship between the number of beds and admissions. The model also fits the data poorly and fails diagnostic tests for model misspecification, meaning that no probative value can be attached to any result that it presents.

In respect of the speciality-specific models, the results offer even less support for the PF's claims regarding the relationship between beds and admissions. Specifically, for the majority (60%) of the specialties considered there is no significant positive relationship between the supply of beds and "discretionary" admissions. The results from the PF's analysis are, therefore, clearly not supportive of its conclusions.

Moreover, there are a number of fundamental shortcomings to the PF's models that serve to undermine the reliability and probative value of the results presented in the PF.

- First, the models are not estimated on the basis of well-defined and meaningful local markets that reflect the range of hospital options individual beneficiaries face. This gives rise to a number of issues that, when addressed, cause the results of the PF's model to change such that there is no longer a significant relationship between the number of beds and admissions.
- Second, the models do not account for the influence of demand on either the level of admissions or the number of beds located within a particular area, an issue that gives rise to endogeneity bias. As a consequence, the positive correlation observed between hospital bed capacity and admissions in the PF's results is not likely to reflect the extent of SID, but rather the fact that hospitals expand capacity in response to unmet demand in local markets. Again, when this issue is accounted for the model results change such that there is no longer a significant relationship between the number of beds and admissions.
- Third, there are a number of explanatory variables included in the PF's models that are likely to be subject to measurement error, meaning that the results presented in the PF are likely to be biased and unreliable.

Fourth, while the PF finds an inverse relationship between local concentration and estimates for "unexplained" increases in expenditure and admissions obtained from its expenditure analysis (i.e. that these estimates are higher in less concentrated local areas), which the PF interprets as evidence of SID, this analysis is again subject to a number of fatal flaws.[3] Indeed, even if the results contained within the PF are taken as given, they do not support the relationship between local concentration and SID.

In particular, the results obtained from the PF's analysis do not support its conclusion for at least two reasons.

- First, the difference in unexplained increases in expenditures between concentrated and non-concentrated local areas is not economically significant, and thus does not provide a

[3] The PF, paragraph 390, page 239.

sound basis to conclude that areas with lower levels of concentration experience higher levels of unexplained increases in expenditure.

- Second, in the case of both expenditures and admissions, the results observed for the intermediate category of moderately concentrated local areas are not consistent with the PF's hypothesis linking local competition to SID, or with the PF's conclusion that facilities located in less concentrated regions have higher levels of expenditure and admissions.

As such, the PF's analysis does not show a systematic or material relationship between the level of local market concentration and differences either between actual and expected admissions rates or between unexplained increases in hospital expenditures.

In addition, the reliability of the inferences that can be drawn from the PF's analysis is fundamentally limited for a number of reasons. The PF's analysis is based on a sample size that amounts to only 12% of the total number of hospitals considered by the PF. Moreover, it appears that the sample of local markets used in the PF's analysis is highly unlikely to be representative of all local markets in South Africa. This strongly suggests not only that the degree of statistical confidence that can be attributed to the results presented in the PF is highly limited, but also that the results are not representative and, therefore, are not reliable.

Moreover, the PF's analysis is based on the results from its cost attribution analysis. For the reasons set out in the RBB report responding to the PHMI's provisional findings on the expenditure analysis, there are a substantial number of reasons to doubt the validity of the results of PF's expenditure analysis.[4] Given that these results are used as inputs in the analysis here, this means that the results of this analysis will also be invalid.

Accordingly, the analysis presented in the PF does not constitute a sound basis to conclude that private hospitals are in any way responsible for any SID that currently exists. This of course also means, in turn, any recommendations targeted at private hospitals with the aim of reducing SID will be ineffective, and thus inappropriate.

Indeed, we also note in any case that the PF's theory is that local areas in which competition between healthcare facilities is more effective (which in its view are areas with lower levels of concentration) are associated with higher levels of SID, and hence lower levels of consumer welfare. This is directly at odds with its finding, in the context of evaluating hospital market power, that "*concentration at the local level […] is therefore a major competitive concern*".[5] Accordingly, any recommendations designed to address former would, by the PF's logic exacerbate the latter, and *vice versa*.

The remainder of the report is structured as follows:

- in Section 1 we discuss the approach taken in the PF to the relationship between facilities, local concentration amongst hospitals and SID;

---

4 RBB, Response to the PHM's Provisional Findings: Expenditure Analysis, 15 October 2018

5 The PF, paragraph 50, page 11.

- in Section 2 we examine the PF's comparison of utilisation levels in South Africa with those of a group of OECD countries;
- in Section 3 we assess the analysis presented in the PF regarding the relationship between local bed capacity and utilisation levels within South Africa; and
- in Section 4 we evaluate the analysis presented in the PF regarding the relationship between local concentration and "unexplained" increases in expenditures and admissions.

1 Facilities, Competition and SID

1.1 Overview

A major conclusion of the PF is that "*healthcare cost escalation is driven by increases in the volume of services performed, rather than increases in the tariffs charged per service*" and finds that "*SID might be one of the causes of increased utilisation of healthcare in the healthcare facilities market in South Africa*".[6]

While the concept of SID has received considerable attention in the economic literature, there is little consensus around how SID should be defined, modelled, or tested.[7] However, properly defining SID, and linking it to a coherent economic model of competition, is necessary if the PF is to reach reliable and robust conclusions around its existence and effects.

In this regard, it is particularly notable that the existing economic literature on SID places healthcare practitioners at the centre, hypothesising that "*medical practitioners have the ability to generate demand for their services directly*".[8] Notably, we are not aware of any economics literature that explicitly assesses the potential role of healthcare facilities in SID. This is principally because healthcare facilities do not act as an agent in the provision of healthcare services and, as a result, cannot act directly to induce demand.

Consequently, it is of crucial importance that the PF, and the models that the PF applies to evaluate SID, are able:

- to provide clear evidence of the mechanism through which facilities (as opposed to practitioners) are able to induce demand; and
- to provide a link between the role of competition between healthcare facilities and SID.

We discuss the approach of the PF in respect of each of these issues below.

1.2 Relationships between facilities and practitioners

The PF appears to acknowledge that it is primarily practitioners that directly induce demand.[9] However, it contends that the contractual relationships between practitioners and hospitals may "*influence practitioners to drive demand*", meaning that hospitals can give rise to SID indirectly.[10] However, the PF's analysis of relationships between practitioners and hospitals does not support this conclusion.

---

6 The PF, paragraph 1, page 376 and paragraph 429, page 245.

7 See, for example, Labelle, R. *et al.* 1994. *A re-examination of the meaning and importance of supplier-induced demand*, Journal of Health Economics, 13(1994), pp. 347-368.

8 Peacock, S. and Richardson, J. 2007. *Supplier-induced demand: re-examining identification and misspecification in cross-sectional analysis*, The European Journal of Health Economics, 8(3), pp. 267-277.

9 See, for example, the PF, paragraph 429.2, page 245 where the PF notes "*practitioners have some discretion round whether to treat and are being paid based on the number of interventions they undertake. This gives them both the ability and incentive for potential manipulation of patients' demand for health services through SID".* Similarly, the PF notes at paragraph 397, page 240 that "*practitioners are directly involved in the clinical diagnosis and the final decision to admit the patient*".

10 The PF, page 217, paragraph 262.

In particular, the contracting arrangements between hospitals and practitioners discussed in the PF do not appear to provide incentives for practitioners to increase the absolute level of admissions. Rather, practitioners that have entered into such arrangements are simply incentivised to ensure that a minimum proportion of their admissions are at the hospital in question.[11]

Such contractual arrangements therefore evidently do not provide practitioners with a perverse incentive to increase overall admissions or utilisation, but instead simply disincentivise practitioners from free riding on relationship specific investments made by facilities. In this regard, the PF itself recognises that "*some alignment of interests between practitioners and facilities would be beneficial, and may make commercial sense … [and] may even promote consumer welfare*".[12]

By the same token, practitioners' shareholdings in healthcare facilities also do not provide a material incentive for practitioners to induce demand. This is illustrated in Life Healthcare's ("LHC's") first submission to the PHMI, which showed that, due to the low level of equity holding, doctors' ability to influence their own returns from a shareholding by inducing demand are materially limited.[13]

This is consistent with the position adopted by the UK Competition and Markets Authority ("CMA"), which considered the competitive effect of practitioner shareholdings in the UK private healthcare market investigation. The CMA concluded that equity participation schemes between private hospital operators and practitioners were not problematic when the equity stake in question was not greater than 5%, subject to certain additional restrictions.[14] **[CONFIDENTIAL]**.[15]

It is, therefore, unlikely that the contractual relationships between practitioners and hospitals, either via contractual obligations or equity shareholdings, have the potential to enable hospitals to indirectly give rise to SID. Indeed, the PF itself notes that it is "*not able to show a direct link between SID and the incentives offered to practitioners*".[16]

1.3 Local competition and SID

Since the objective of the PHMI is to find whether there "*is any feature or combination of features of markets in the private health care sector which harm competition or has an adverse effect on competition within that market*", it also necessary for the PF to link the potential existence of SID to local competition between healthcare facilities.[17]

The PF proposes that SID arises as a result of the two-stage nature of competition between healthcare facilities. Specifically, according to the PF, competition takes place:

11 The PF, paragraph 256, page 216.
12 The PF, paragraph 261, page 218.
13 LHC's First Submission to the health Market Inquiry, dated 31 October 2014, paragraph 14.3.3.
14 CMA, "Private healthcare market investigation", Final Report, 2 April 2014, paragraph 11.461.
15 LHC's First Submission to the health Market Inquiry, dated 31 October 2014, paragraph 14.3.3.13.
16 The PF, paragraph 261, page 217.
17 The PF, paragraph 3, page 14.

1. at a national level, where facilities and funders negotiate nationally for tariffs and network inclusion; and
2. then at a local level, were facilities compete on the basis of patient volumes by attracting practitioners.

The PF proposes that hospitals' role in SID arises via the second-stage local level competition. Specifically, the PF proposes that healthcare facilities do not compete directly for patient volumes at a local level, but rather compete to attract and retain the "*best admitting doctors*", who will thus, in turn, attract patients to that facility.[18]

The PF suggests a number of ways through which this 'competition for practitioners' takes place. These include investing in unnecessary equipment, high bed capacity, hiring a high number of nurses, renting consultation rooms to practitioners at below market rates, and/or offering shareholding to doctors. Once private healthcare facilities have attracted a doctor, the PF finds that the contracts entered into between practitioners and facilities may then provide incentives for practitioners to increase utilisation. Accordingly, since practitioners have some discretion around treatment, the PF finds that this gives practitioners the "*ability and incentive for potential manipulation of patients' demand for health services through SID*".[19]

The implication of the PF's theory is that local areas in which competition between healthcare facilities is more effective (i.e. areas with lower levels of concentration) are associated with higher levels of SID, and hence lower levels of consumer welfare, than areas in which local competition between healthcare facilities is less effective (i.e. areas with high levels of concentration). Specifically, the PF hypothesises that:

> *"More local competition … leads to higher systemic costs and inefficiencies in terms of more admissions, overcapacity, lower utilisation, overtreatment and higher expenditure on matters that generally are beneficial to the doctors, but not necessarily to the patient – and by implication not to the payer/member of a scheme".*[20]
>
> *"The overall result may be counter-intuitive in terms of the economic orthodoxy: less concentrated and competitive markets may imply inefficient, higher than necessary average costs of treatment and vice versa".*[21]

Notably, the PF's theory of SID is therefore directly at odds with its finding, in the context of evaluating hospital market power, that "*concentration at the local level […] is therefore a major competitive concern*".[22] Indeed, if one were to accept the PF's hypothesis that more local competition results in SID and overutilisation of healthcare services, this would imply that less

18 The PF, paragraph 122, page 185.
19 The PF, paragraph 429.2, page 245.
20 The PF, paragraph 286, page 138.
21 The PF, paragraph 380, page 237-238.
22 The PF, paragraph 50, page 11.

local competition (i.e. higher levels of local concentration) would in fact be preferable to more local competition (i.e. low levels of local concentration).

This results in a fundamental tension between the PF's conclusions regarding SID and its conclusions around local concentration and the effectiveness of competition between private healthcare facilities. It also creates a fundamental tension with the PF's recommendations aimed at reducing levels of local concentration through, *inter alia*, forced divestiture and specifying maximum allowable levels of local concentration.[23] This is also discussed further in the RBB report responding to the PHMI's provisional findings on the effectiveness of competition.[24]

---

23 The PF, paragraph 78-81, page 465.

24 RBB, Response to the PHM's Provisional Findings: Effectiveness of Competition, 15 October 2018.

2 Healthcare Utilisation Levels

2.1 Overview

In this section we discuss the PF's cross-country comparison of healthcare utilisation levels. We start by briefly summarising the PF's analysis and results, before discussing the appropriateness of the comparator countries included in the PF's analysis. We then comment on the adequacy of the PF's attempt to account for demand-side factors that are likely to drive cross-country differences in healthcare utilisation rates. Lastly, we consider the results obtained from the PF's assessment and evaluate whether these support the PF's conclusions.

2.2 The PF's assessment of healthcare utilisation levels

As a starting point in its assessment of SID, the PF compares hospital utilisation rates in South Africa for total hospital admissions and selected "discretionary" procedures with those of 17 OECD comparator countries. The PF's particular focus on "discretionary" procedures is adopted on the basis that these procedures are the most likely candidates for SID, since it is in respect of these procedures that "*practitioners have some discretion around whether to treat".*[25]

Further, the PF's motivation for comparing utilisation rates in South Africa to those in OECD countries is that utilisation rates in these countries "*should represent a 'high water mark' for demand unconstrained by resources".*[26] In other words, the PF appears to contend that these countries are likely to provide a benchmark for levels of utilisation that are neither affected SID (and thus inflated) or by resource constraints (and thus depressed).

Following its cross-country comparison of utilisation rates, the PF finds that:

- overall hospital utilisation rates were higher in South Africa than in all but 2 OECD comparator countries, after adjusting for differences in the age distribution of country's populations;
- age-adjusted admissions rates were higher in South Africa than the average of all comparator countries for 6 out of 7 "discretionary" procedures; and
- ICU admissions rates were higher in South Africa than in 8 comparator countries.

On this basis the PF concludes that for "*selected diagnosis where there is discretion around whether or not to admit a patient, South Africa tends to over-service compared to OECD countries*".[27] In other words, the PF considers these findings to be evidence of the presence of SID in South Africa.

[25] The PF, paragraph 4, page 377.
[26] The PF, paragraph 9, page 378.
[27] The PF, paragraph 429.5, page 245.

2.3 Evaluation of comparator countries

In respect of the OECD comparator countries selected, the PF notes that "*all of the comparator countries citizens have universal coverage through publicly funded national health or insurance schemes [sic]*" and that since all of the comparator countries also "*have significantly higher GDP that South Africa it was felt that utilisation rates in each should represent a 'high water mark' for demand unconstrained by resources [sic]*".[28]

However, it is important to appreciate that in order for observed differences in utilisation rates between South Africa and the OECD comparator countries to be due to the presence of SID in South Africa, it is important to control for other factors that could cause utilisation rates across countries to differ. An obvious example of this is where different countries face different resourcing constraints, such that in more constrained countries utilisation rates are lower simply because there is not the capacity in the healthcare system to serve demand.

In an apparent attempt to address this issue, the PF asserts that the OECD comparator countries have "*unconstrained resources".*[29] However, they provide no support for this assertion and, in fact, the evidence available to us suggest that a large number of the comparator countries are subject to significant resource constraints in the healthcare sector.

In particular, the extent of average waiting times provides an obvious metric through which to evaluate the extent of resource constraints for healthcare procedures, and in this regard the OECD notes that:

> *"[l]ong waiting times for healthcare services is an important policy issue in many OECD countries … although less relevant in some (e.g. Belgium, France, Germany, Japan, Korea, Luxembourg, Switzerland, United States)".*[30]

The linkage between waiting times and resourcing constraints, and thus in turn utilisation, is borne out by a number of pieces of evidence.

- First, it appears that the comparator countries with the lowest hospital utilisation rates are also the comparator countries with the longest waiting times for healthcare services. For instance, the 8 OECD comparator countries with the lowest hospital utilisation rates have all been found to have long waiting times.[31] Likewise, the two comparator countries with the lowest utilisation rates - Portugal and Spain - have been found to have among the longest waiting times for elective surgery of the OECD countries – 61 days for Spain and 86 days for Portugal (compared to 33 days for France).[32]

---

28 The PF, paragraph 9, page 378.

29 The PF, paragraph 9, page 378.

30 OECD. 2017. "Waiting Times for elective surgery", in *Health at a Glance 2017: OECD Indicators,* page 96. Available: https://doi.org/10.1787/health_glance-2017-28-en

31 Siciliani, L. and J. Hurst, Explaining Waiting-Time Variations For Elective Surgery Across OECD Countries, O.P. OECD Health Working Papers No 7 and d. 10.1787/406746186162, Editors. 2003, OECD.

32 2010 Commonwealth Fund International Health Policy Survey in Eleven Countries

- Second, the OECD comparator countries with the highest hospital utilisation rates are those with the shortest waiting times for healthcare facilities. Indeed, the two comparator countries with admissions rates higher than that of South Africa, namely Germany and Austria, are widely recognised as having very short waiting times.[33]

Therefore, not only does it appear that resource constraints are likely to be a significant determinant of differences in the hospital utilisation rates observed across countries, but it also appears that a large proportion of the comparator counties with lower utilisation rates than South Africa have significant resource constraints, while those with higher utilisation rates do not suffer from such constraints.

The implication of this is that the OCED comparator group of countries simply does not provide a reliable benchmark for utilisation rates in the absence of SID. Hence, any comparisons made between South Africa and this comparator group are invalid, and do not provide evidence of the existence of SID in South Africa. Indeed, if anything, South African utilisation rates ought to be compared to those comparator countries least likely to be subject to resource constraints, such as Germany.

2.4 Age-adjustment

The PF correctly acknowledges that, in order to assess whether SID is a driver of utilisation of private healthcare services in South Africa, one must not only control for other supply-side factors such as resource constraints (see sub-section 2.3), but also demand-side factors. Accordingly, the PF states that they have sought to compare the utilisation rates for discretionary procedures to those in the OECD comparator countries only "*after adjusting for acceptable demand drivers such as age, illness prevalence, and severity, as well as other insurance market failures, such as adverse selection*".[34]

However, and notwithstanding this list of factors, in practice the only demand-side factor that the PF seeks to control for is differences in the age profile of South Africa and the OECD comparator countries. This is significant since, as the PF appears to acknowledge, there are many other risk factors that may drive differences in admissions rates, and if these are not controlled for this will serve to further invalidate the cross-country comparison.

For example, age distribution is not the only, or even the primary, demand side factor that is likely to result in different utilisation rates for caesarean sections between counties. More important is likely to be the proportion of the population that is female, and the fertility rate. Similarly, in the case of CABG procedures, lifestyle factors are likely to be an important driver of utilisation rates, in addition to the age distribution.

In addition, while the PF acknowledges that it is necessary to account for adverse selection, no attempt has been made to do so. As noted in the Cadiant report on the PHMI's provisional findings ("the Cadiant report"), South African medical schemes are likely to suffer from adverse

[33] 2010 Commonwealth Fund International Health Policy Survey in Eleven Countries and Viberg, N. *et al.* 2013. *International comparisons of waiting times in healthcare – Limitations and prospects.* Health Policy 112 (2013), pp 53-61.

[34] The PF, paragraph 5, page 377.

selection, and, as a result, South African beneficiaries are likely to be less healthy, all else equal, than the population of countries with universal healthcare coverage.[35]

In short, therefore, the cross-country comparison presented in the PF does not seek to control for a variety of demand-side factors. This means that utilisation rates across countries are not comparable, and thus provides a further reason why the PF's cross-country analysis of utilisation rates has no probative value for assessing the existence, or extent, of SID in South Africa.

2.5 Comparison of utilisation rates

Notwithstanding the fatal deficiencies of the comparison made between South Africa and the OECD comparator group countries, as well as the failure to account for relevant demand-side factors, the PF's comparison of utilisation rates does not, in any case, conclusively show that South African rates exceed those of the OECD comparator countries. Moreover, if errors in the PF's calculations are corrected for and South African utilisation rates are compared to the most appropriate comparator country, namely Germany, the PF's conclusion that "*South Africa tends to over-service"* is not supported.[36]

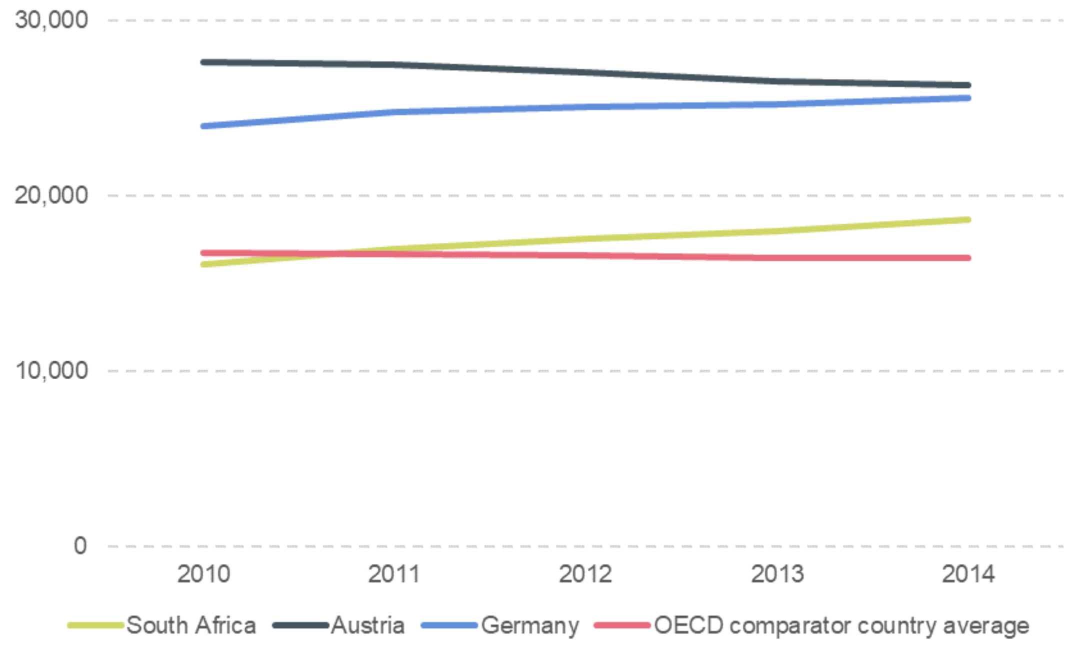
2.5.1 Overall hospital admissions

As noted in the Cadiant report, the PF's methodology in adjusting for the presence of day-cases in the South Africa admissions data appears to have been incorrectly and inconsistently applied. In particular, the PF has not correctly accounted for the presence of day admissions in the South Africa figures, and thereby has overstated the age-adjusted admissions rates for South Africa.[37]

In this regard, the corrected age-adjusted admissions rates for South Africa presented in the Cadiant report show that admissions rates in South Africa are in line with those of the OECD comparator countries. Moreover, age-adjusted overall hospital admissions rates in South Africa are materially lower than those in Germany and Austria. This is apparent from Figure 1 below, which shows the age-adjusted hospital admissions rates for South Africa, Austria and Germany, as well as the average rate across all 17 OECD countries considered in the PF's analysis, over the period 2010 to 2014.

35 Cadiant Partners, HMI Draft Findings and Recommendations, September 2018.
36 The PF, paragraph 429, page 246.
37 Cadiant report, page 8.

Figure 1: Comparison of age-adjusted hospital admissions rates for South Africa and select comparator countries



Source: PHMI hospital admissions data and age-adjusted hospital admissions rates presented in the Cadiant report, page 8.

2.5.2 Procedure-level comparison

The PF's procedure-level comparison of utilisation rates between South Africa and the OECD comparator countries is also not conclusive. In particular, as shown in Table 1 below, age-adjusted utilisation rates in South Africa are lower or not materially different from the average OECD comparator country rate for 4 out of the 7 "discretionary" procedures considered in the PF's analysis.[38]

This suggests that, when compared to the OECD comparator country average, there is no evidence that South Africa over-services in respect of the majority of "discretionary" procedures considered by the PF. Moreover, when compared to Germany, age-adjusted admissions rates in South Africa are higher in only 3 out of 7 "discretionary" procedures.[39]

---

[38] Specifically, age-adjusted utilisation rates in South Africa are lower or not materially different from the average OECD comparator country rate for inguinal hernia, cholecystectomy, CABG, and major joint arthroplasty procedures.

[39] Specifically, age-adjusted utilisation rates in South Africa are higher than that of Germany for cataract surgery, tonsillectomy, and caesarean section procedures.

Table 1: Comparison of procedure-level age-adjusted utilisation rates for South Africa and select OECD comparator countries, 2010 to 2014

Procedure	South Africa	Germany	OECD comparator country average
Inguinal hernia	1.4	**2.2**	1.9
Cataract surgery	**12.5**	9.6	8.3
Cholecystectomy	1.9	**2.4**	1.8
CABG	0.6	**0.7**	0.4
Major joint arthroplasty	3.4	**4.9**	3.4
Tonsillectomy	**2.6**	1.9	1.3
Caesarean section	**6.9**	3.1	2.5

Source: *PHMI age-adjusted admissions rates.*
Notes: *Admissions rates for "Major joint arthroplasty" procedures have been corrected for comparability as discussed in the Cadiant report, page 11.*

Consequently, the PF's procedure-level cross-country comparison of utilisation rates does not provide evidence that age-adjusted utilisation rates for "discretionary" procedures are systematically higher in South Africa than in comparator OECD countries.

Moreover, as noted in LHC's response to the PF, the level of discretion that doctors have in respect of the procedures analysed in the PF is likely to be limited due to the risks associated with not treating a patient that displays symptoms that are typically treated by such procedures. In particular, LHC notes that:

- cholecystectomy procedures treat the cause of chronic and agonising abdominal pain, which often affects women;
- tonsillectomy procedures treat one of the leading causes of Upper Airways Obstruction in repeated throat infections children, which, if untreated, can lead to secondary complications that may affect the heart;
- major joint arthroplasty treats arthritis of major joints which is the leading cause of chronic debilitating pain in patients above 65 years of age;
- inguinal hernia repair procedures, while not an emergency procedure, can lead to life threatening strangulation if not performed;
- cataract surgery treats one of the leading causes of blindness in South Africa and is necessary to restore someone's vision; and
- coronary artery bypass grafting for coronary ischemia is a life-saving complex surgery for patients with Ischaemic Heart Disease.[40]

As a result, it is not clear that the "discretionary" procedures included in the PF's analysis are, in fact, truly procedures in respect of which "*practitioners have some discretion around whether*

40 LHC Response to the PF, dated 15 October 2018, paragraph 3.12.6.

to treat".[41] This serves to undermine the probative value of PF's procedure-level analysis is in assessing the likely extent of SID in South Africa further.

2.6 Summary

The PF's comparison of hospital utilisation rates in South Africa for total hospital admissions and selected "discretionary" procedures with those of 17 OECD comparator countries (which the PF contends are likely to provide a benchmark for levels of utilisation that are neither affected SID or resource constraints) are subject to a number of fundamental shortcomings.

Most notably, the PF's analysis does not, in any case, show that South African rates systematically exceed those of the OECD comparator countries. In particular, overall hospital age-adjusted admissions in South Africa are only marginally higher than the OECD comparator country average rate, while age-adjusted utilisation rates in South Africa are lower or not materially different from the average OECD comparator country rate for 4 out of the 7 "discretionary" procedures considered in the PF's analysis.

Moreover, it appears that resource constraints are likely to be a significant determinant of differences in the hospital utilisation rates observed across countries. In this regard, it is notable that a large proportion of the comparator countries with lower utilisation rates than South Africa have significant resource constraints, while those with higher utilisation rates than South Africa do not suffer from such constraints. Consequently, the age-adjusted utilisation rates are not readily comparable for the purposes of assessing the potential existence of SID. Moreover, South Africa's utilisation rates tend to be similar or lower than those in countries that appear likely to be unconstrained in terms of healthcare resources.

In addition, it is not clear that the "discretionary" procedures included in the PF's analysis are in fact truly discretionary procedures, and the PF fails to account for a number of demand-side factors that are likely to drive cross-country differences in utilisation rates. This ultimately means that the utilisation rates that have been compared are not readily comparable. This provides a further reason why any differences between countries cannot be reliably ascribed to the existence of SID.

As a result, the PF's cross-country analysis of utilisation rates does not provide any robust support for the existence of SID in South Africa, or the PF's conclusion that "*South Africa tends to over-service".*[42]

41 The PF, paragraph 4, page 377.
42 The PF, paragraph 429, page 246.

3 Relationship between Capacity and Healthcare Utilisation

3.1 Overview

In this section we discuss the PF's assessment of the relationship between healthcare capacity and utilisation. We start by summarising the methodology adopted by the PF to examine this relationship, before considering whether the results presented in the PF, if taken at face value, support the PF's conclusions regarding SID. We then discuss whether the PF's model is valid and likely to yield reliable results, focussing in particular on the geographic level at which the model is estimated, the impact of endogeneity, and the likelihood that the model's explanatory variables are subject to measurement error.

3.2 The PF's assessment of the relationship between healthcare capacity and utilisation

In an attempt to determine the likely existence and extent of SID in South Africa, the PF conducts a regression analysis that seeks to ascertain whether, consistent with its theory of SID, there exists a positive relationship between the supply of practitioners and/or hospital beds and levels of utilisation, when all other relevant explanatory factors are controlled for.[43] If such a relationship were identified (and found to be material), this would suggest that practitioners and/or healthcare facilities are driving increased healthcare expenditure by inducing beneficiaries to over use private healthcare services.

The regression analysis is conducted using beneficiary data obtained from medical schemes, together with hospital claims data, covering the period 2010 to 2014. These data are combined with data reflecting the estimated number of hospital beds and doctors within 234 municipal areas throughout South Africa.[44]

The PF estimates a series of logistic regression models that attempt to model the likelihood that a beneficiary is admitted to hospital under a particular specialty in a given year. The likelihood of admission is modelled as a function of patient characteristics, as well as the supply of doctors and beds in the municipality in which a given beneficiary resides.[45]

The PF estimates models for overall hospital admissions, as well as for admissions under eleven specialities, which the PF classifies as "*specialities with high rates of discretionary admission*".[46] In respect of the specialty-specific regression models, the PF first models the likelihood that a beneficiary is admitted under that specialty before modelling the likelihood that the beneficiary is admitted for a procedure that the PF has defined as "discretionary".[47]

---

[43] The PF, paragraph 17, page 382.
[44] The PF, paragraph 7, page 377 and paragraph 18, page 382-383.
[45] The patient characteristics included in the PF's model include age, level of medical scheme coverage, chronic illnesses, year of treatment, and the amount of time that a beneficiary has been a member of the medical scheme in an attempt to capture the effect of adverse selection
[46] Specifically, the PF selected orthopaedics, ENT surgery, neurosurgery, psychiatry, urology, cardiology, cardio-thoracic surgery, paediatrics, general surgery, obstetrics, and general surgery.
[47] The PF, paragraph 27, page 387.

The PF finds that the per capita supply of hospital beds are a positive and significant predictor of overall hospital admissions and, similarly, that the number of doctors operating in a municipal area is a significant positive driver of overall hospital admissions. In respect of discretionary admissions for the eleven specialities considered, the PF finds that the supply of beds is a significant positive predictor of admissions for four out of eleven specialties, while the supply of doctors is a positive significant predictor of discretionary admissions for nine out of eleven specialties.

On the basis of these results, the PF concludes that the "*study shows evidence that rates of hospital admission are positively associated with levels of supply of hospital beds, after adjusting for clinical and demographic factors. The greater the proportion of hospital beds to the local population, the higher the rate of admissions in a given region*".[48] Further, the PF finds that this provides evidence in support of SID operating via private hospital bed capacity (i.e. that bed capacity leads to SID).[49]

3.3 The results from the PF's analysis do not support its conclusion

From the outset, it must be appreciated that, contrary to the PF's claims, its conclusions are not supported by the results from their analysis, particularly in respect of the relationship between admissions rates and hospital bed capacity. In particular, in order to show that private healthcare facilities engage in SID one would need to find evidence that the supply of hospital beds systematically gives rise to materially higher admissions rates.

However, as far as the PF's overall admissions model is concerned, while the estimated relationship between the supply of hospital beds and the likelihood of a beneficiary being admitted to hospital is found to be statistically significant (i.e. differentiable from zero with a reasonable degree of statistical confidence) the size of the estimated relationship is not material. This is illustrated by the estimated marginal effects for the PF's model presented in Table 2 below.[50] [51]

In particular, the mean marginal effects estimate for beds implies that an increase of 1 bed per 100 beneficiaries in a particular municipal area increases the probability of beneficiary being admitted to hospital by only 0.8%. In contrast, an increase of 1 doctor per 100 beneficiaries increases the probability of a beneficiary being admitted to hospital by 26%. These results therefore show that there at best a very weak relationship between the number of beds and admissions.

48 The PF, paragraph 19, page 383.

49 The PF, paragraph 34, page 390.

50 Marginal effects can be described as the change in outcome as a function of the change in the treatment (or independent variable of interest) holding all other variables in the model constant. These are easier to interpret from a practical standpoint than the log odds ratios presented in the PF.

51 We were not able to run regressions and additional statistical tests using the full dataset employed in the PF analysis due to constraints on the processing power of computers provided by the PHMI. However, PHMI has confirmed that the smaller data sample provided is representative of the full dataset and, as a result, our additional analysis should yield similar results when estimated over the full dataset. Indeed, the results presented in Annex A below show that when replicating the PF's overall admissions model we obtain materially identical results to those presented in the PF.

Table 2: Selected marginal effects estimates for the PF's overall admissions model

Explanatory variable	Mean marginal effects
Beds	0.008***
	(0.001)
Doctors	0.260***
	(0.014)
Pseudo-R^2	0.084

Source: *RBB analysis of PHMI data.*
Notes: ** reflects statistical significance at the 10% level, ** reflects statistical significance at the 5% level, and *** reflects statistical significance at the 1% level.*

Moreover, the PF's overall admissions model fits the data poorly and, even more importantly, it fails standard diagnostic tests for model misspecification. As shown Table 2 above, the pseudo-R^2 reported for the model indicates that the model is only able to explain 8.4% of the observed variation in the probability that a beneficiary is admitted to hospital at any point in the year. This suggests that the model has very little explanatory power.

By the same token, the results from the "link" test, presented in Annex B below, show that the PF's model likely suffers from misspecification.[52] This indicates that it is likely that the model fails to include all relevant factors that affect the likelihood of a beneficiary being admitted to hospital and/or has not specified these relationships correctly.

Unless all relevant explanatory variables are included in the model, the estimates derived from it will be biased and unreliable (i.e. the model will not accurately reflect true relationship between the various explanatory variables and admissions). In turn, this means that no reliable or robust inferences can be drawn from the model's estimates.

In respect of the speciality-specific models presented in the PF, the results are even less supportive of the PF's conclusions, since they fail to show any systematic positive relationship between the supply of beds and the likelihood of a beneficiary being admitted for a discretionary procedure. In particular:

- for only <u>4 out of 10 specialties</u> considered did the PF find a <u>positive</u> and statistically significant relationship between the supply of hospital beds and "discretionary" admissions;
- for <u>2 out of 10 specialties</u> considered the PF finds a <u>negative</u> and statistically significant relationship between the supply of hospital beds and "discretionary" admissions; and
- for <u>4 out of 10 specialities</u> considered the PF finds <u>no</u> statistically significant relationship between the supply of hospital beds and "discretionary" admissions.

Therefore, in respect of the majority (60%) of the specialties considered, the PF finds no significant positive relationship between the supply of beds and "discretionary" admissions.

[52] Pregibon, D. 1980. Goodness of link tests for generalized linear models. Applied Statistics 29: 15–24.

Similarly, when considering overall speciality-specific admissions (as opposed to "discretionary" admissions), the PF finds no significant positive relationship between the supply of beds and admissions for 70% of the specialities considered.

Therefore, notwithstanding the limitations of the PF's models set out below in sub-sections 3.4 and 3.5, the PF's analysis in any event fails to support its conclusion that "*rates of hospital admissions are positively associated with levels of supply of hospital beds*".[53]

3.4 The PF's models are not estimated using meaningful local markets

The PF attempts to model the probability that a beneficiary is admitted to hospital based on the number of hospital beds (and supply of doctors) in the local market in which that beneficiary resides. The PF's hypothesis is that, due to SID, all else equal, a beneficiary in a local market with more beds will be more likely to be admitted to hospital.

This implies that the PF should be estimating the relationship between the total number of beds at hospitals that a given beneficiary would consider attending (i.e. the total number of beds in the local market of a given beneficiary) and the likelihood that that beneficiary is admitted to hospital in a given year. However, if the model is estimated based on the total number of beds at a set of hospitals that differs from the set of hospitals that the beneficiary in question actually considers as viable options, the estimated relationships will have no economic or practical meaning, and thus no probative value in the assessment of the extent of SID.

In this regard, it is notable that the PF's model is not estimated at the level of well-defined local markets for healthcare facilities that reflect the range of options individual beneficiaries face. Instead, it is estimated at the level of municipal areas (i.e. the PF estimates the relationship between the number of beds in the municipality in which a beneficiary resides and the probability that beneficiary will be admitted to hospital). However, municipal areas do not constitute appropriately defined local markets, and are not derived on the basis of the private healthcare facilities that are available to patients. For the reasons explained above, this means that the PF's analysis cannot have any probative value in assessing SID.

Moreover, this issue is not merely a theoretical one, but is, in practice, extensive. In particular, approximately 33% of beneficiary records are for beneficiaries who live in a municipality that has zero hospital bed capacity. The assumption of the PF's model is that, since there are no hospitals available for that beneficiary to choose to visit, none of these beneficiaries should be admitted to hospital. However, 76% of the beneficiaries who live in a municipality that has zero hospital bed capacity were in fact admitted to hospital.[54]

Even more importantly, this has a material effect on results presented in the PF. In particular, as shown in Table 3 below, adjusting the PF's overall admission model to include a variable that accounts for the effect of beneficiaries that live in areas with no hospital beds gives rise to results (in the PF's model) that show that the supply of hospital beds has no statistically

[53] The PF, paragraph 427, page 245.
[54] RBB analysis of PHMI data. The calculation is based on a representative sample of the full dataset used in the PF.

significant effect on admissions at the 5% confidence level.[55] Moreover, accounting for the effect of these beneficiaries results in the estimated marginal effect being more than halved.

As shown in Annex C, this result is robust to running the regression over only the sample of beneficiaries that live in areas with beds present. Moreover, while we have not presented results from similar extensions to the PF's speciality-specific models, we would expect that similar conclusions apply to these models.

Table 3: Comparison of mean marginal effects for PF's model and a model accounting for beneficiaries in municipalities with no hospital beds

	Original model	Accounting for no beds
Beds	0.008***	0.003*
	(0.001)	(0.001)
Doctors	0.260***	0.130***
	(0.014)	(0.020)
Dummy variable	-	-0.021***
	-	(0.002)

Source: RBB analysis of PHMI data.
*Notes: * reflects statistical significance at the 10% level, ** reflects statistical significance at the 5% level, and *** reflects statistical significance at the 1% level.*

Therefore, as a consequence of failing to estimate the models on the basis of well-defined and meaningful local markets (that reflect the range of hospital options individual beneficiaries face), the PF's results hold no probative value for the assessment of the existence of SID in respect of private healthcare facilities. Indeed, as shown above, if one accounts for the effect of this shortcoming hospital bed capacity is not found to be a statistically significant driver of hospital admissions.

3.5 The PF's models suffer from endogeneity

A further limitation of the PF's analysis is that it does not account for the influence of demand on either the level of admissions or the number of beds located within a particular area. This leads to an endogeneity problem, meaning that the results of the PF's models will be biased and, therefore, unreliable. In turn, this means that the results from the PF's analysis, even assuming no other problems exist with the model, will not constitute reliable evidence of SID.

The omission of relevant explanatory variables, i.e. variables that are able to explain some part of the outcome of interest, leads to bias in any regression model. Moreover, if an omitted explanatory variable is related to even one of the included explanatory variables, all of the estimated results from the regression model will generally be biased and, hence, unreliable.[56]

55 Specifically, we include a simple dummy variable equal to one when a beneficiary resides in a municipal area with no hospital beds present and equal to zero otherwise.
56 Wooldridge, J. (2000). *Introductory Econometrics: A Modern Approach* (First Edition). United States of America: Thomson Learning, p.91.

This problem is broadly referred to as endogeneity, since both the dependent variable and the explanatory variable in question are endogenously determined.

In the case at hand, endogeneity arises since areas in which there are more hospital beds are more likely to be areas in which demand for hospital services is higher, since bed capacity is likely to have been expanded to meet such demand. As a result, areas with higher demand for hospital services will also have higher admissions rates. However, a positive correlation between bed capacity and admissions in this situation may simply reflect that hospital groups build facilities in the areas with greatest demand for their services, rather than reflecting SID. Importantly, unless this endogeneity is controlled for, it is not possible to isolate any potential effect arising due to SID from the effect of demand.

A robust way of doing so would be to control for endogeneity through an instrumental variables ("IV") regression.[57] However, due to limitations of the data available to us we have not been able to estimate an IV regression. As a result, we have adopted an approach, which, while unlikely to address the issue of endogeneity fully, is likely to reduce its extent, and therefore illustrate the impact that endogeneity plays in driving the results presented in the PF.

In particular, we have re-run the PF's overall regression model including dummy variables to account for municipal fixed effects. In effect, this accounts for variation, including in the level of demand, from one municipal area to the next, and thereby partially addresses the problem associated with the PF's model not accounting for the effect that demand has on admission rates.[58]

As shown below in Table 4, when the PF's model is expanded to include municipal fixed effects, the estimated relationship between the supply of doctors and hospital admissions is no longer statistically significant. While we have not presented similar results for the PF's speciality-specific models, we would expect that including municipal fixed effects in these models would have similar impacts on the results of those models.

Table 4: Comparison of log odds ratio for PF's model and a model including municipal fixed effects

	Original model	Municipal fixed effects
Beds	1.06***	1.14
	(0.010)	(0.099)
Doctors	6.67***	11.74***
	(0.691)	(2.12)

Source: RBB analysis of PHMI data.
*Notes: * reflects statistical significance at the 10% level, ** reflects statistical significance at the 5% level, and *** reflects statistical significance at the 1% level.*

This provides a strong indication that the results presented in the PF are largely driven by endogeneity that arises from a failure to account for the impact that demand has on both

57 See, for example, Wooldridge, J. M. 2001. *Econometric analysis of cross section and panel data*. MIT Press, Chapter 5.

58 See, for example, Wooldridge, J. M. 2001. *Econometric analysis of cross section and panel data*. MIT Press, Chapter 10.

hospital capacity and admissions. As a consequence, the positive correlation observed between hospital bed capacity and admissions in the PF's results is not likely to reflect the extent of SID, but rather the fact that hospitals expand capacity in response to unmet demand in local markets.

3.6 The PF's models are subject to measurement error

Notwithstanding the difficulties with the PF's models discussed above in sub-sections 3.4 and 3.5, there are a number of issues with the underlying data that limit the reliability and probative value of the PF's results. It is important to recognise that the validity and reliability of the results obtained from any statistical analysis will be constrained by quality and relevance of the underlying data used to perform such analysis.[59] Principle considerations in this regard include whether the data accurately measure the explanatory variables and outcome(s) of interest, whether the data are consistently measured and compiled throughout the sample, and whether the data are themselves recorded accurately.

In this regard, there are a number of factors which suggest that the quality of the data used to estimate the PF's models is limited. Most notably, there are a number of reasons to doubt the accuracy of the data reflecting the distribution of beds across municipalities. Indeed, the PF's have estimated the distribution of beds by "*consolidating and interpolating bed numbers for some hospitals and years*" and conclude that, while imperfect, the data capture the "*macroscopic distribution of private beds across the country*"".[60]

However, the PF's models are estimated at the municipal level and as a result, since the data reflect only the aggregate distribution of beds across the country, the explanatory variable capturing the supply of hospital beds in a given municipality is likely to be subject to measurement error. In this regard, it is notable that where there is measurement error amongst any of the explanatory variables, the model will suffer from bias, thereby rendering its estimates inaccurate and unreliable.[61]

Moreover, as noted in the Cadiant report, the PF has used the "narrow disease burden" of patients to account for the existence of chronic or pre-existing medical conditions.[62] For the reasons set out in the RBB Economics report responding to the PHMI's provisional findings on the expenditure analysis and the Cadiant report, this variable is likely to be a poor measure of disease burden and, as such, is likely to be subject to measurement error. [63] Similarly, the Cadiant report notes that there are a number of issues with the PF's categorization of scheme plans, which mean that this variable is likely to be a poor measure of scheme coverage and, therefore, also subject to measurement error.

---

59 The terms reliability and validity have specific meanings in statistical and regression analysis. Reliability refers to the repeatability of the findings from the analysis (i.e. would the results be similar to those obtained from a similar analysis), while validity refers to the credibility of the findings (i.e. did the analysis measure what it was intended to measure?).

60 The PF, paragraph 65, page 401.

61 Wooldridge, J. M. 2001. *Econometric analysis of cross section and panel data*. MIT Press, page 51.

62 Cadiant report, page 11.

63 RBB, Response to the PHM's Provisional Findings: Expenditure Analysis, 15 October 2018.

Consequently, there are a number of explanatory variables included in the PF's models that are likely to be subject to measurement error. This means that the results presented in the PF are likely to be biased and unreliable.

3.7 Summary

In summary, the PF's regression analysis, which seeks to determine whether there exists a positive relationship between the supply of practitioners and/or hospital beds and levels of utilisation (after controlling for relevant explanatory factors) does not support its conclusion that healthcare facilities are inducing beneficiaries to overuse healthcare services through SID.

Specifically, the results from the PF's model of overall hospital admissions show that there is, at best, a very weak relationship between the number of beds and admissions. In any case, the model fits the data poorly and fails diagnostic tests for model misspecification, so that no probative value can be attached to any result that it presents.

In addition, in respect of the speciality-specific models, for the majority (60%) of the specialties considered there is no significant positive relationship between the supply of beds and "discretionary" admissions. This result is clearly at odds with the PF's conclusion.

Moreover, there are a number of fundamental shortcomings to the PF's models that serve to undermine the reliability and probative value of the results presented in the PF.

First, the models are not estimated on the basis of well-defined and meaningful local markets (that reflect the range of hospital options individual beneficiaries face). This gives rise to a number of issues that, when corrected for mean that the PF's model shows that there is no significant relationship between the number of beds and admissions.

Second, the models do not account for the influence of demand on either the level of admissions or the number of beds located within a particular area. As a consequence, the positive correlation observed between hospital bed capacity and admissions in the PF's results is not likely to reflect the extent of SID, but rather the fact that hospitals expand capacity in response to unmet demand in local markets. Again, when this issue is accounted for in the PF's model, there is not significant relationship between the number of beds and admissions.

Third, there are a number of explanatory variables included in the PF's models that are likely to be subject to measurement error, meaning that the results presented in the PF are likely to be biased and unreliable.

Overall therefore, the PF's analysis does not provide any credible support for its conclusion that "*rates of hospital admissions are positively associated with levels of supply of hospital beds"*.[64]

[64] The PF, paragraph 427, page 245.

4 Relationship between Concentration and SID

4.1 Overview

In this section we consider the PF's analysis of the relationship between local market concentration and SID. We first briefly summarise the PF's analysis, and then consider whether the results from the PF's analysis, if taken at face value, actually support the PF's conclusions. Finally, we set out a number of shortcoming to the PF's analysis, which indicate that its results are in any case likely to be unreliable.

4.2 The PF's analysis of concentration and SID

In seeking to assess the relationship between concentration of private healthcare facilities and SID, the PF compares levels of "unexplained" expenditure and admissions in concentrated markets with that in non-concentrated local markets. In doing so, the PF assesses whether differences in the level of local competition translates into differences in expenditure and admissions between local markets as a result of SID.

As discussed above in section 1, the PF's hypothesise that more competitive local markets may result in higher costs, as facilities compete to attract practitioners beyond what is efficient (a hypothesis that we note would motivate for intervention in the opposite direction to the more conventional competition intervention that is also proposed by the PF).

In its assessment, the PF categorises 24 selected local markets into concentrated, moderately concentrated, and non-concentrated markets, according to the HHI (although these definitions appear to us somewhat arbitrary for the reasons discussed in more detail in the RBB report responding to the PHMI's provisional findings on the facilities concentration analysis).[65] Two analyses are then performed:

- First, the PF considers the difference between actual and expected admission rates to assess whether admission patterns are different in markets with different levels of concentration. The PF hypothesises that, as a consequence of SID, non-concentrated local markets should exhibit larger differences between actual and expected admissions rates than concentrated local markets.
- Second, total and unexplained expenditure increases are assessed by concentration level as per the cost attribution analysis. The PF's SID hypothesis implies that unexplained expenditure increases should be higher in non-concentrated local markets than concentrated local markets.

The PF finds that concentrated local markets have lower admission rates and lower cost increases than non-concentrated regions. Although it does acknowledge that this evidence is by no means conclusive, the PF takes this finding as support for the view that *"facilities located in concentrated regions display characteristics which could be considered perverse"*, and

[65] RBB, Response to the PHM's Provisional Findings: Facilities Concentration Analysis, 15 October 2018.

considers this perversity to indicate that *"hospital groups at the national level have the market power to compensate for inefficient competition at the local level"*.[66]

4.3 The results of the PF's analysis do not support its conclusion

From the outset, it is notable that, notwithstanding the flaws with the analysis discussed below, the results obtained from the PF's analysis do not support its conclusion for at least two reasons.

First, there is no material difference in unexplained increases in expenditures between concentrated and non-concentrated local markets. In particular, as shown in Table 5 below, in 2014 there was a difference of 0.1 percentage point in unexplained expenditure increases between concentrated and non-concentrated local markets. This difference is not economically significant, and does not provide a meaningful basis on which to conclude that regions with lower levels of concentration experience higher levels of unexplained increases in expenditure

Second, in the case of both expenditures and admissions, the results observed for moderately concentrated local markets are not consistent with the PF's hypothesis linking local competition to SID, or with the PF's conclusion that facilities located in less concentrated regions have higher levels of expenditure and admissions. In particular, as shown in Table 5 below, in respect of both expenditure and admissions, moderately concentrated areas are found to have higher levels than both concentrated and non-concentrated regions. This suggests that there is no systematic relationship between the level of local concentration (competition) and the level of unexplained expenditure increases, contrary to the PF's SID hypothesis.

Table 5: Summary of the PF's analysis of admissions and expenditure across regions, 2014

Region	Admissions (difference from expected value)	Expenditure (unexplained increase)
Concentrated	-8.4%	2.3%
Moderately concentrated	11.3%	3.5%
Non-concentrated	7.2%	2.4%

Source: PF Table 6.13 and Table 6.14

As a result, the PF's analysis of levels of "unexplained" expenditure and admissions across local markets with differing levels of concentration yields results that are not consistent with its SID hypothesis. In particular, the PF's analysis shows that there is no systematic relationship or material relationship between the level of local market concentration and either differences between actual and expected admissions rates or unexplained increases in hospital expenditures.

66 The PF, paragraph 390, page 239.

4.4 The PF's analysis is subject to a number of additional shortcomings

The reliability of the inferences that can be drawn from any statistical analysis will, to a large extent, depend on the degree to which the data sample used contains information related to the actual economic outcomes, processes, and relationships of interest. In this regard, the smaller the data sample size the less useful information the data contain relating to the actual outcomes and processes of interest. Moreover, the smaller the sample size the more likely it will be that information contained in the data sample that is unrelated to the economic relationships of interest (often referred to as "noise") will affect the results of the analysis. [67]

In this regard, the small sample size used in the PF's analysis raises significant concerns about the reliability of the results it presents. Specifically, the PF considers a sample of only 24 hospitals, which amounts to only 12% of the total number of hospitals considered by the PF in the local concentration analysis. This strongly suggests that the degree of statistical confidence that can be attributed to the results presented in the PF is likely to be highly limited.

In addition to sample size, the reliability of results from any statistical analysis will depend on the degree to which the data sample used is representative of the complete population from which the sample has been drawn. In other words, if the sample of local markets used in the PF's analysis is not representative of all local markets in South Africa the results from the analysis will not be reliable. In this regard, while the PF does not provide details on how the sample was selected, there is reason to expect that the sample used in the PF's analysis is not representative since the "*sample regions do not therefore partition the entire country or the entire membership base used".*[68] This suggests that the results from the PF's analysis are not representative and, therefore, are not reliable.

Lastly, it is notable that the PF's analysis is based on the results from its cost attribution analysis. For the reasons set out in the RBB report responding to the PHMI's provisional findings on the expenditure analysis, there are a substantial number of reasons to doubt the validity of the PF's expenditure analysis.[69] As a consequence, this casts significant doubt over the validity of the PF's analysis of levels of "unexplained" expenditure and admissions across local markets with differing levels of concentration.

---

67 Davidson and MacKinnon. 2004. *Econometric Theory and Methods.* Oxford University Press. Page 93.

68 NMG Consultants and Actuaries, *Expenditure analysis report 4: Facilities analysis,* page 48.

69 RBB, Response to the PHM's Provisional Findings: Expenditure Analysis, 15 October 2018

Annexes

A Comparison of overall regression results

Table 6: Comparison of the PF's and RBB's coefficient estimates for the overall admissions model

	PF estimates	RBB estimates
Beds	0.029***	0.060***
	(0.001)	(0.010)
Doctors	0.070***	1.89***
	(0.011)	(0.103)
Psuedo-R^2	0.080	0.084

Source: *RBB analysis of PHMI data.*
Notes: ** reflects statistical significance at the 10% level, ** reflects statistical significance at the 5% level, and *** reflects statistical significance at the 1% level.*

B Results from the link test

Table 7: Parameter estimates from link test on the PF's overall admissions model

Parameter	Estimate
Prediction	0.793***
	(0.013)
Prediction squared	-0.079***
	(0.005)
Constant	-0.098***
	(0.010)
Log likelihood	-216081.28
LR chi2(2)	39795.80
Prob > chi2	0.0000
Pseudo-R^2	0.0843

Source: *RBB analysis of PHMI data.*
Notes: ** reflects statistical significance at the 10% level, ** reflects statistical significance at the 5% level, and *** reflects statistical significance at the 1% level.*

C Model estimated over sample of beneficiaries residing in municipal areas with bed capacity

Table 8: Comparison of mean marginal effects for PF's model and a model accounting for beneficiaries in municipalities with no hospital beds

	Original model	Dummy variable for no bed capacity	Excluding beneficiaries with no bed capacity
Beds	0.008***	0.003*	0.003*
	(0.001)	(0.001)	(0.002)
Doctors	0.260***	0.130***	0.167***
	(0.014)	(0.020)	(0.021)
Dummy variable	-	-0.021***	-
	-	(0.002)	-

Source: RBB analysis of PHMI data.
*Notes: * reflects statistical significance at the 10% level, ** reflects statistical significance at the 5% level, and *** reflects statistical significance at the 1% level.*